일반논문 (Regular Paper)
방송공학회논문지 제22권 제3호, 2017년 5월 (JBE Vol. 22, No. 3, May 2017)
https://doi.org/10.5909/JBE.2017.22.3.366
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

DCT를 통한 UWV 콘텐츠의 2D 인접도 행렬 생성

이 소 율[a)], 김 규 헌[a)‡]

2D Adjacency Matrix Generation using DCT for UWV Contents

Li Xiaorui[a)] and Kyuheon Kim[a)‡]

요 약

TV나 디지털 사이니지와 같은 화면표시장치들이 점점 커져감에 따라, 미디어의 종류가 UHD, 파노라마, 퍼즐형 미디어와 같은 광각의 미디어로 변하고 있다. 특히, 파노라마 및 퍼즐형 미디어는 스티칭을 통해 복수개의 카메라로 촬영된 비디오 클립을 합성한 형태로 구성된다. 그러나, 스티칭 과정의 처리 시간이 오래 걸리기 때문에 실시간 서비스에는 적용하기 어려운 문제가 있다. 따라서 본 논문에서는 스티칠 처리 시간을 감소하기 위한 방법으로, 영상간의 공간적 연관관계를 알려주는 2D Adjacency Matrix를 생성하는 것을 제안한다. Discrete Cosine Transform (DCT)를 사용하여, 비디오 소스의 각 프레임을 공간 영역에서 주파수 영역으로 변환 시킨다. 앞서 언급한 DCT 계수를 기반으로 효과적으로 이미지들의 공간적 연관관계를 알려주는 2D Adjacency Matrix를 생성한다. 본 논문에서는 각각의 비디오 클립들로부터 파노라마 영상과, 퍼즐형 미디어를 생성하기 위해 DCT를 이용한 2D Adjacency matrix 생성 방법을 제안한다.

Abstract

Since a display device such as TV or digital signage is getting larger, the types of media is getting changed into wider view one such as UHD, panoramic and jigsaw-like media. Especially, panoramic and jigsaw-like media is realized by stitching video clips, which are captured by different camera or devices. However, a stitching process takes long time, and has difficulties in applying for a real-time process. Thus, this paper suggests to find out 2D Adjacency Matrix, which tells spatial relationships among those video clips in order to decrease a stitching processing time. Using the Discrete Cosine Transform (DCT), we convert the each frame of video source from the spatial domain (2D) into frequency domain. Based on the aforementioned features, 2D Adjacency Matrix of images could be found that we can efficiently make the spatial map of the images by using DCT. This paper proposes a new method of generating 2D adjacency matrix by using DCT for producing a panoramic and jigsaw-like media through various individual video clips.

Keyword : 2D Adjacency Matrix, Discrete Cosine Transform (DCT), panorama

a) 경희대학교 전자정보대학(Department of Electronic Engineering, Kyung Hee University)
‡ Corresponding Author : 김규헌(Kyuheon Kim)
E-mail: kyuheonkim@khu.ac.kr
Tel: +82-31-201-3810
ORCID: http://orcid.org/0000-0003-1553-936X
※ This research was partly supported by the ICT R&D program of MSIP/IITP (B0126-16-1013, Development of Generation and Consumption of Jigsaw-like Ultra-Wide Viewing Spatial Media).
· Manuscript received December 19, 2016; Revised February 28, 2017; Accepted March 31, 2017.

Ⅰ. Introduction

There are various methods to increase the immersion experience for users, such as the author improve video reality and immersion by matching between odor and color[1]. However, expanding the filed of view is the most direct and effective way. Images captured by a single device, such as a cell-phone, a DV or a tablet, often have a limited field of view. By combing several overlapped images captured by multiple cameras[2], a view expanded stitching image can be obtained, which is often referred to panorama[3] or image mosaics[4]. In recent years, with the development of the display technology, the use of the images shows an increasing tendency. Thus, the size of the image, which is exhibited in the display, is also expected to be enlarged. However, the view of normal camera is only 60 degree which makes the difficulty in conveying the realistic feeling to users in some vast areas, such as stage and stadium. To solve this problem, the method, which uses several cameras to generate a large image from multiple angle and multiple viewpoint, is proposed.

There exist lots of algorithms, taking the SIFT (Scale-invariant feature transform) as a most frequently used method for example, which takes the advantages of transforms image data into scale-invariant coordinates relative to local features[5] to composite images. Using feature points to combine the image is benefit to the synthesis part, since it can calculate the exact combination position and is also robust for image distortion. However, due to the mass of compared feature points and processing information, the computation is relatively large along with the low speed. As a result, a speeded-up algorithm named SURF (Speeded up robust features) is proposed, which is partly inspired by the SIFT descriptor. The authors claims that the standard version of SURF is several times faster than SIFT and more robust against different image transformations than SIFT[6]. However, in the case of stitching many images, it still can't meet the requirements of fast stitching of multiple images. Kang, Seung Heon, et al[7]. parallelize SIFT and SURF, and its application to fast panoramic image generation on the latest embedded GPU. This paper proposes the use of a 2D Adjacency Matrix, which tells the similar region among images to be stitched, and a stitching process is only applied to the region defined by a 2D Adjacency Matrix and is resulted in shorter stitching processing time with a similar stitching quality.

The discrete cosine transform (DCT)[8] used in image compression methods such as JPEG and MPEG[9], which is a variant of the Fourier Transform suitable for many image processing applications. Using DCT-domian approach to retain the image information can not only reduce the complexity of computation but also supplying better accuracy[10]. Choi, Sung-Woo, and Oh-Seol Kwon[11] utilize the energy compaction quality to compensate the illuminant in the DCT-domain. In this paper we propose 2D Adjacency Matrix Generation which utilize the DCT block similarity detection method inspecting the similarity of the input images and then obtaining the image map. TThe rough spatial relationship of input images, which can be utilized for further image stitching, is generated through the image map. On the basis of this approach, a sticking process will be much simpler with less computation time. Thus, a user could watch a wider view contents, which is generated from a sticking process with multiple video sequences captured by different cameras. This wider view contents will bring more immersive effect to a user over a conversion platform of both a broadcasting and a communication channels.

In present work, several parts are involved, namely, the description of the DCT matching contains DCT algorithm and DCT block similarity detection method (section 2), the 2D Adjacency Matrix Generation (see section 3), and section 4 presents the experimental results and comparison are contained. Finally, the conclusion of the proposed system and the view of future work on this topic are stated in section 5.

Ⅱ. DCT Matching

1. DCT algorithm

The DCT (Discrete Cosine Transform), and in particular the DCT-II, is often used in signal and image processing, especially for lossy compression, because it has a strong "energy compaction" property[12, 13].

DCT-II

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right] \qquad k = 0, \ldots, N-1.$$

그림 1. DCT-II 계산식
Fig. 1 . The DCT-II calculating formula

2. DCT Block Similarity Detection Method

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{\bar{H}_1 \bar{H}_2 N^2}} \sum_I \sqrt{H_1(I) H_2(I)}}$$

그림 2. 바타차리야 거리 계산식
Fig. 2. The Bhattacharyya distance calculating formula

The transformed DCT images of the similar images possess semblable data distribution, while the DCT images of different images have another data distribution. On account of this point, DCT comparison may be an approach for detecting the similarity of images. Regard to DCT comparison, it consists of several steps. First of all, the comparing region is divided into many 8 x 8 blocks and DCT transform is applied on each block. Secondly to filter the pixels which value is small than 1, a 8 x 8 Zigzag matrix is utilized for doing this work. And then an equation in Fig.2 named Bhattacharyya distance[14] is used to calculate the distance of comparing blocks from the DCT coefficient of each transformed block. Finally we get the average value of distance of the all divided 8 x 8 blocks as a standard for further analysis.

In terms of the DCT block similarity detection method, it can be divided into several steps. In brief, two images are firstly input. Subsequently, the DCT comparison is carried out to the whole of the input images. When the calculated result exceeds the specified threshold 0.15, it can be concluded that these two input images are very similar as well as possess the same position. Otherwise, these images are divided into two parts in the horizontal and vertical direction, respectively. To further determine the relation between two input images, the DCT comparison and analysis on the segmented blocks are performed. Start here the threshold standard for judging the similarity of comparing parts will change to the minimum value of DCT comparison.

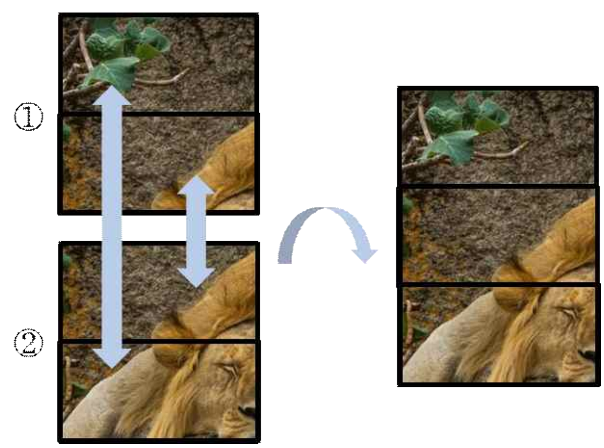


그림 3. 수직 방향 세분화된 이미지와 매칭 결과
Fig. 3. Segmented images in the vertical direction and matching result

The comparison of lower part of the image 1 and the upper portion of the image 2 using DCT block similarity detection method is executed, as demonstrated in Fig. 3. If they were judged as similar ones, it is reasonable to confirm that the image 2 is below the image 1 and shows 1/2 overlap with the image 1. On the other hand, if the upper portion of the image 1 was similar to the lower part of the image 2, one obtains that the image 1 is on the up position of the image 2.

Moreover, by utilizing the DCT block similarity detection method, the similarity result between the right portion of the image 1 and the left portion of the image 2 is

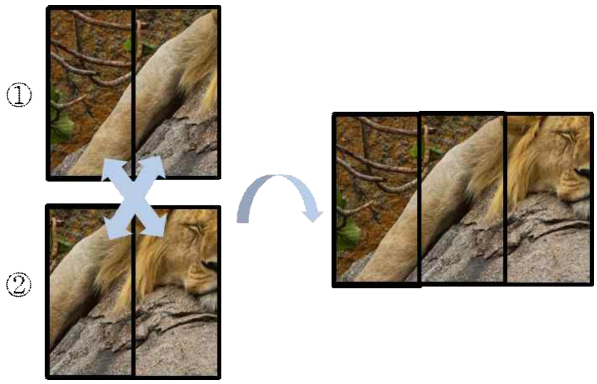


그림 4. 수직 방향 세분화된 이미지와 매칭 결과
Fig. 4. Segmented images in the horizontal direction and matching result

achieved, as shown in Fig. 4. If the similarity was high, it indicates that the image 1 is situated at the left side of the image 2 along with 1/2 overlap with the image 2. In another case, if the left portion of the image 1 and right portion the image of 2 was similar, one knows that the image 1 is located at the right side of the image 2.

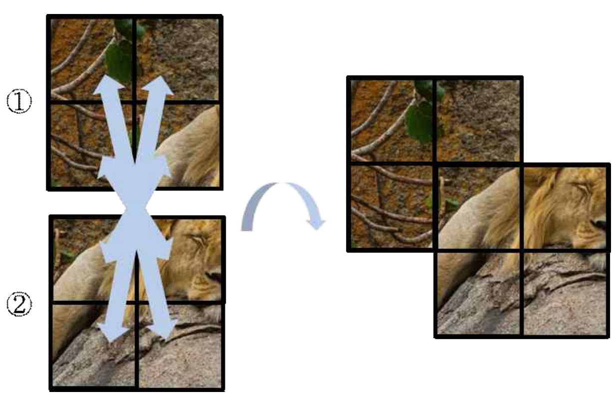


그림 5. 대각선 방향 세분화된 이미지 매칭 결과
Fig. 5. Segmented images in the diagonal direction and matching result

If the similarity of the input images can not be achieved through aforementioned procedure, the image should be divided into 4 parts for further analysis. With the help of DCT comparison, we can know the position relation of segmented images in the diagonal direction. As being shown in Fig. 5, the comparison result between the lower right part of the image 1 and the upper left part of the image 2 by DCT comparison, where it is evident that the image 1 is placed at upper left side of the image 2 with the overlap of 1/4 with the image 2 in the diagonal direction.

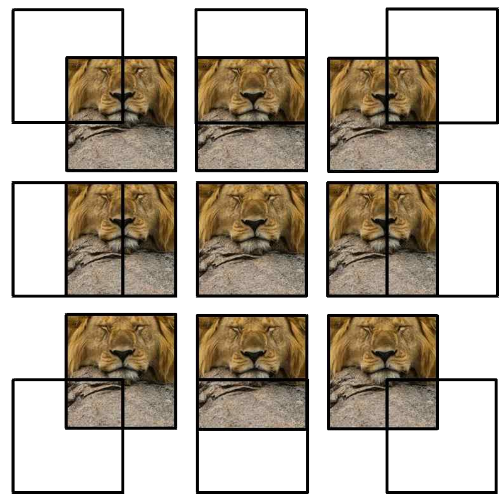


그림 6. 사용 가능 영역
Fig. 6. Possible locations

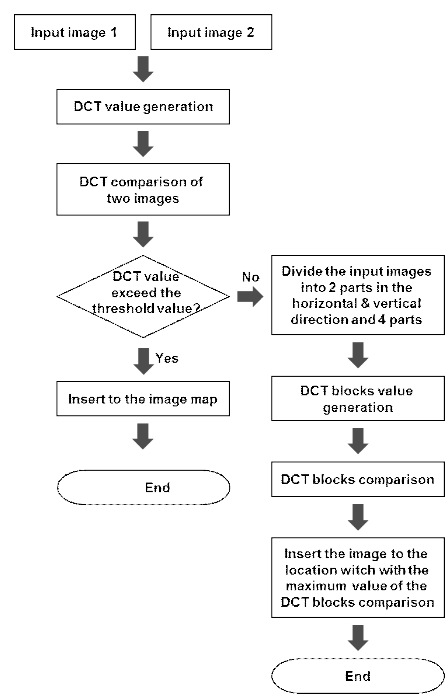


그림 7. 2D 인접도 행렬 생성 알고리즘 흐름도
Fig. 7. 2D Adjacency Matrix Generation algorithm flow chat

When two images are compared with each other, the potential confirmed positions are investigated and corresponding results are depicted in Fig. 6. Clearly, eight positions are obtained from an image (see Fig. 6). The schematic diagram for the above algorithm is displayed in Fig. 7.

Ⅲ. 2D Adjacency Matrix Generation

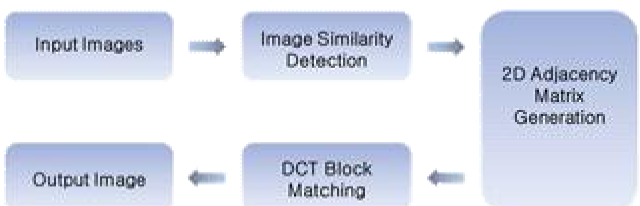


그림 8. 시스템 블락 도표
Fig. 8. System block diagram

The operation procedure of the proposed system, which is used in this work, is demonstrated in Fig. 8. In this case, the grayscale images with the resolution of 256 x 256 are firstly inserted. Then, the DCT comparison is implemented to inspect the similarity of the input images. On the basis of the similarity of the input images, one knows their 2D Adjacency Matrix. Ultimately, by means of DCT block matching process on the input images, the image map would be created.

The image position can be gained via the DCT block similarity detection method. Based on the values obtained from DCT comparison, the expression of 2D Adjacency Matrix between the images is achieved and it could be employed to create the image map. As is known, the image map is a collection of the input images. In general case, if the image A was shot at the right side of the image B, it would be populated at the right side of the image B in the image map.

If we already had an image map and want to do further image mosaic, such as mosaicking panorama and jigsaw-like image, the amount of calculation would be greatly reduced, leading to the enormous enhancement of computing speed.

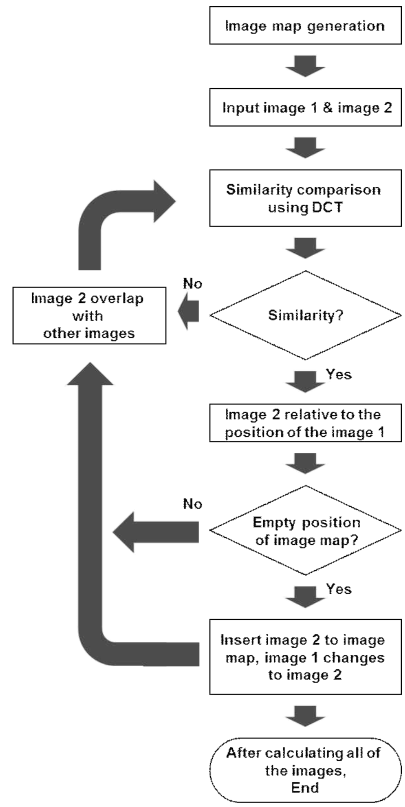


그림 9. 이미지 지도 생성 알고리즘 흐름도
Fig. 9. Image map generation algorithm flow chat

To produce the image map, the images are initially input into the system. After inputting the images, the DCT comparison method is employed to analyze the similarity among the image 1 with other input images. In particular, when the image 1 is located at a certain position, from the similarity result, the position of the image 2, which is related to that of the image 1, could be determined. After finishing the comparison between the image 1 and other input images, the images, which show the similarity with the image 1, will be applied as the reference image to continually compare with other residual images. The detailed processes are illustrated in Fig. 9.

In terms of generation of an image map, the most critical thing is that each image can only occupy a certain position in the image map, that is to say, different images possess different positons. If two images were very similar to each other, one of them should be removed. To realize this purpose, an alignment process is required, as described in Fig. 9. In this case, the image 2 is expected to inset into the image map. Unfortunately, the candidate inserting position, which is related to the image 1, is occupied by another image. To figure it out, the insertion of the image 2 into the image map is quitted and it is used to detect the similarity with other images.

Ⅳ. Experimental Results and Comparison

To verify the performance of the proposed algorithm, two different sets of images are tested as the experimental object with the resolution of 256x256 and 512x512, respectively. The experimental environment consists of Microsoft Visual Studio 2013 from Microsoft Company, OpenCV2.4.10 library[15] and 3.4GHz Intel I7 processor.

1. Image map test sets and mapping result

1.1. First test image set and mapping result

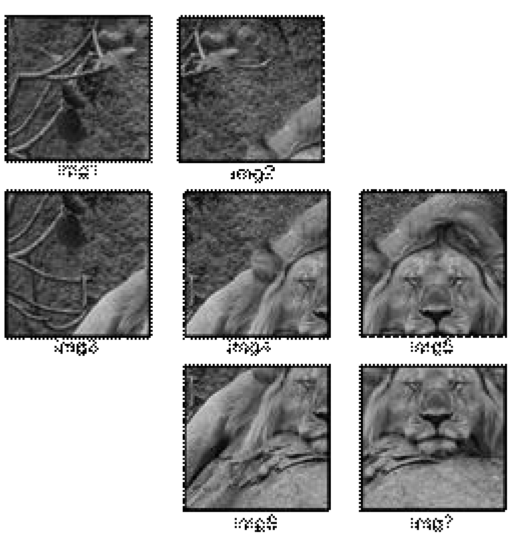


그림 10. 첫 번째 테스트 이미지
Fig. 10. First test images

그림 11. 첫 번째 테스트 이미지의 이미지 지도
Fig. 11. Image map of first test images

The first test image set consists of 7 images with 256x 256 resolution are shown in Fig. 10. The mapping image of the input images is generated, as shown in Fig. 11. The runtime of the proposed system for generation of the image map is 0.107s.

1.2. Second test image set and mapping result

The second test image set consists of 18 images with 256x256 resolution respectively are shown in Fig. 12. Eventually, the image map of the input images is generated, as depicted in Fig. 13. The mapping time of producing the map image is 8.692s.

그림 12. 두 번째 테스트 이미지
Fig. 12. Second test images

그림 13. 두 번째 테스트 이미지의 이미지 지도
Fig. 13. Image map of second test images

2. Comparison run-time of different algorithm

그림 14. 입력 이미지
Fig. 14. Input images

In this part, the stitching results of 4 input images are presented, witch adopted proposed mapping method with ROI SURF stitching and pure SURF stitching respectively. On the other hand, the run-time of two algorithms are measured and shown in Table I and Fig. 14 is the input images with 512x512 resolution.

In Fig. 15. (a), it can be seen that the input images are well aligned by utilizing 2D Adjacency Matrix Generation using DCT method as illustrated previously. Then referring to the spatial relationship of the input images obtained before, ROI SURF algorithm is adopted to stitch the aligned images to a stitching image, as depicted in Fig. 15. (b). Moreover, the performance of pure SURF stitching result can be confirm in Fig. 16.

Table I shows the execution time of proposed method with ROI SUFR stitching and pure SURF stitching algorithms for execution. The classical SURF stitching without mapping process takes 18 times longer than the proposed method in this paper as being shown in Table I.

(a)

(b)

그림 15. 2D 인접도 행렬 이미지 맵의 성능과 ROI SURF 스티칭 결과 (a) 2D 인접도 행렬 이미지 맵, (b) ROI SURF 스티칭 결과
Fig. 15. Performance of 2D Adjacency Matrix image map with ROI SURF stitching result: (a) 2D Adjacency Matrix image map, (b) ROI SURF stitching result

그림 16. 순수 SURF 스티칭 결과
Fig. 16. Performance of pure SURF stitching result.

표 1. 다른 알고리즘의 실행시간 비교
Table 1. Comparison of run-time of different algorithms

Algorithm	Generation time for image map (Sec.)	Total run-time (Sec.)
SURF Stitching	\	59.224s
2D Adjacency Matrix image map generation with ROI SURF stitching	0.619s	3.163s

Ⅴ. Conclusion and Future Work

In conclusion, based on the experimental results, it is evident that DCT block similarity detection method is suitable for detecting the similarity between different images. Meanwhile, the proposed method is also an effective approach to deal with these slightly warping images such as a shorter runtime as low as 3.163s for stitching 4 images. However, as for processing the rotated and larger warping images, there still exist some problems. In the future work it is needed to solve the problem for improving the matching accuracy. As a consequence, the goal of the future work is to overcome these drawbacks.

참 고 문 헌 (References)

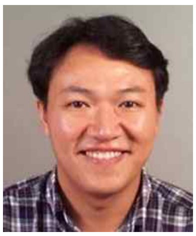
[1] Lee, Guk-Hee, et al. "The Effect of Matching between Odor and Color on Video Reality and Sense of Immersion." Journal of Broadcast Engineering 19.6 (2014): 877-895.
[2] Guo, Heng, et al. "Joint Video Stitching and Stabilization From Moving Cameras." IEEE Transactions on Image Processing 25.11 (2016): 5491.
[3] Szeliski, Richard, and Heung-Yeung Shum. "Creating full view panoramic image mosaics and environment maps." Proceedings of the 24th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co., 1997.
[4] Shum, H-Y., and Richard Szeliski. "Construction of panoramic image mosaics with global and local alignment." Panoramic vision. Springer New York, 2001. 227-268.
[5] Lowe, David G. "Distinctive image features from scale-invariant keypoints."International journal of computer vision 60.2 (2004): 91-110.
[6] Speeded up robust features (https://en.wikipedia.org/wiki/Speeded_up_robust_features)
[7] Kang, Seung Heon, et al. "Parallelization of Feature Detection and Panorama Image Generation using OpenCL and Embedded GPU." Journal of Broadcast Engineering 19.3 (2014): 316-328.
[8] Chen, Wen-Hsiung, and C. Smith. "Adaptive coding of monochrome and color images." IEEE Transactions on Communications 25.11 (1977): 1285-1292.
[9] Le Gall, Didier. "MPEG: A video compression standard for multimedia applications." Communications of the ACM 34.4 (1991): 46-58.
[10] Pandey, Achala, and Umesh C. Pati. "A novel technique for mosaicing of medical images." 2014 Annual IEEE India Conference (INDICON). IEEE, 2014.
[11] Choi, Sung-Woo, and Oh-Seol Kwon. "Face Recognition using High-order Local Pattern Descriptor and DCT-based Illuminant Compensation." Journal of Broadcast Engineering 21.1 (2016): 51-59.
[12] Narasimha, M., and A. Peterson. "On the computation of the discrete cosine transform." IEEE Transactions on Communications 26.6 (1978): 934-936.
[13] Makhoul, John. "A fast cosine transform in one and two dimensions." IEEE Transactions on Acoustics, Speech, and Signal Processing 28.1 (1980): 27-34.
[14] Bhattacharyya, A. "On a Measure of Divergence between Two Multinomial Populations." Sankhyā: The Indian Journal of Statistics (1933-1960), vol. 7, no. 4, 1946, pp. 401–406., www.jstor.org/stable/25047882.
[15] OpenCV Reference (http:// docs.opencv.org)

저 자 소 개

이 소 율

- 2014년 2월 : 중국 HEBEI UNIVERSITY 자동화공학 학사
- 2017년 2월 : 경희대학교 전자전파공학 석사
- ORCID : http://orcid.org/0000-0002-0920-8882
- 주관심분야 : 영상처리, 멀티미디어 시스템

저 자 소 개



김 규 헌

- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006년 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 교수
- ORCID : http://orcid.org/0000-0003-1553-936X
- 주관심분야 : 영상처리, 멀티미디어통신, 디지털 대화형방송